



RAW SEQUENCE LISTING ERROR REPORT

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) detected errors when processing the following computer readable form:

Application Serial Number: 10/653,681
Source: 1FWO
Date Processed by STIC: 3/29/04

THE ATTACHED PRINTOUT EXPLAINS DETECTED ERRORS.

PLEASE FORWARD THIS INFORMATION TO THE APPLICANT BY EITHER:

- 1) INCLUDING A COPY OF THIS PRINTOUT IN YOUR NEXT COMMUNICATION TO THE APPLICANT, WITH A NOTICE TO COMPLY or,
- 2) TELEPHONING APPLICANT AND FAXING A COPY OF THIS PRINTOUT, WITH A NOTICE TO COMPLY

FOR CRF SUBMISSION AND PATENTIN SOFTWARE QUESTIONS, PLEASE CONTACT MARK SPENCER, TELEPHONE: 703-308-4212; FAX: 703-308-4221

Effective 12/13/03: TELEPHONE: 571-272-2510; FAX: 571-273-0221

TO REDUCE ERRORED SEQUENCE LISTINGS, PLEASE USE THE CHECKER VERSION 4.1 PROGRAM, ACCESSIBLE THROUGH THE U.S. PATENT AND TRADEMARK OFFICE WEBSITE. SEE BELOW FOR ADDRESS.

<http://www.uspto.gov/web/offices/pac/checker/chkr4lnote.htm>

Applicants submitting genetic sequence information electronically on diskette or CD-Rom should be aware that there is a possibility that the disk/CD-Rom may have been affected by treatment given to all incoming mail.

Please consider using alternate methods of submission for the disk/CD-Rom or replacement disk/CD-Rom.

Any reply including a sequence listing in electronic form should NOT be sent to the 20231 zip code address for the United States Patent and Trademark Office, and instead should be sent via the following to the indicated addresses:

1. EFS-Bio (<<http://www.uspto.gov/ebc/efs/downloads/documents.htm>> , EFS Submission User Manual - cPAVE)
2. U.S. Postal Service: Commissioner for Patents, P.O. Box 1450, Alexandria, VA 22313-1450
3. Hand Carry directly to (EFFECTIVE 12/01/03):
U.S. Patent and Trademark Office, Box Sequence, Customer Window, Lobby, Room 1B03, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202
4. Federal Express, United Parcel Service, or other delivery service to: U.S. Patent and Trademark Office, Box Sequence, Room 1B03-Mailroom, Crystal Plaza Two, 2011 South Clark Place, Arlington, VA 22202

Revised 10/08/03

Raw Sequence Listing Error Summary

ERROR DETECTED	SUGGESTED CORRECTION	SERIAL NUMBER: <u>10/653,681</u>
ATTN: NEW RULES CASES: PLEASE DISREGARD ENGLISH "ALPHA" HEADERS, WHICH WERE INSERTED BY PTO SOFTWARE		
1 _____ Wrapped Nucleics Wrapped Aminos	The number/text at the end of each line "wrapped" down to the next line. This may occur if your file was retrieved in a word processor after creating it. Please adjust your right margin to .3; this will prevent "wrapping."	
2 _____ Invalid Line Length	The rules require that a line not exceed 72 characters in length. This includes white spaces.	
3 _____ Misaligned Amino Numbering	The numbering under each 5 th amino acid is misaligned. Do not use tab codes between numbers; use space characters , instead.	
4 _____ Non-ASCII	The submitted file was not saved in ASCII(DOS) text, as required by the Sequence Rules. Please ensure your subsequent submission is saved in ASCII text.	
5 _____ Variable Length	Sequence(s) _____ contain n's or Xaa's representing more than one residue. Per Sequence Rules, each n or Xaa can only represent a single residue. Please present the maximum number of each residue having variable length and indicate in the <220>-<223> section that some may be missing.	
6 _____ PatentIn 2.0 "bug"	A "bug" in PatentIn version 2.0 has caused the <220>-<223> section to be missing from amino acid sequences(s) _____. Normally, PatentIn would automatically generate this section from the previously coded nucleic acid sequence. Please manually copy the relevant <220>-<223> section to the subsequent amino acid sequence. This applies to the mandatory <220>-<223> sections for Artificial or Unknown sequences.	
7 _____ Skipped Sequences (OLD RULES)	Sequence(s) _____ missing. If intentional, please insert the following lines for each skipped sequence: (2) INFORMATION FOR SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) (i) SEQUENCE CHARACTERISTICS: (Do not insert any subheadings under this heading) (xi) SEQUENCE DESCRIPTION:SEQ ID NO:X: (insert SEQ ID NO where "X" is shown) This sequence is intentionally skipped Please also adjust the "(ii) NUMBER OF SEQUENCES:" response to include the skipped sequences.	
8 _____ Skipped Sequences (NEW RULES)	Sequence(s) _____ missing. If intentional , please insert the following lines for each skipped sequence. <210> sequence id number <400> sequence id number 000	
9 _____ Use of n's or Xaa's (NEW RULES)	Use of n's and/or Xaa's have been detected in the Sequence Listing. Per 1.823 of Sequence Rules, use of <220>-<223> is MANDATORY if n's or Xaa's are present. In <220> to <223> section, please explain location of n or Xaa, and which residue n or Xaa represents.	
10 _____ Invalid <213> Response	Per 1.823 of Sequence Rules, the only valid <213> responses are: Unknown, Artificial Sequence, or scientific name (Genus/species). <220>-<223> section is required when <213> response is Unknown or is Artificial Sequence	
11 _____ Use of <220>	Sequence(s) _____ missing the <220> "Feature" and associated numeric identifiers and responses. Use of <220> to <223> is MANDATORY if <213> "Organism" response is "Artificial Sequence" or "Unknown." Please explain source of genetic material in <220> to <223> section. (See "Federal Register," 06/01/1998, Vol. 63, No. 104, pp. 29631-32) (Sec. 1.823 of Sequence Rules)	
12 _____ PatentIn 2.0 "bug"	Please do not use "Copy to Disk" function of PatentIn version 2.0. This causes a corrupted file, resulting in missing mandatory numeric identifiers and responses (as indicated on raw sequence listing). Instead, please use "File Manager" or any other manual means to copy file to floppy disk.	
13 _____ Misuse of n/Xaa	"n" can only represent a single <u>nucleotide</u> ; "Xaa" can only represent a single <u>amino acid</u>	



IFWO

RAW SEQUENCE LISTING

DATE: 03/29/2004

PATENT APPLICATION: US/10/653,681

TIME: 13:18:10

Input Set : A:\U 014798-3.ST25.txt

Output Set: N:\CRF4\03292004\J653681.raw

3 <110> APPLICANT: Dai, Ken-Shwo
 5 <120> TITLE OF INVENTION: HUMAN ARL-RELATED GENE VARIANTS ASSOCIATED WITH CANCER
 7 <130> FILE REFERENCE: U 014798-3
 9 <140> CURRENT APPLICATION NUMBER: 10/653,681
 10 <141> CURRENT FILING DATE: 2003-09-02
 12 <160> NUMBER OF SEQ ID NOS: 4
 14 <170> SOFTWARE: PatentIn version 3.2
 16 <210> SEQ ID NO: 1
 17 <211> LENGTH: 1090
 18 <212> TYPE: DNA
 19 <213> ORGANISM: ARTIFICIAL
 21 <220> FEATURE:
 22 <223> OTHER INFORMATION: UNKNOWN
 25 <220> FEATURE:
 26 <221> NAME/KEY: CDS
 27 <222> LOCATION: (70)..(333)
 29 <400> SEQUENCE: 1
 30 caaaaacagc aacagaaagc aggacgtgag acttctacct gctcactcag aatcatttct
 32 gcaccaacc atg gcc acg ttt gtg gag ctc agt acc aaa gcc aag atg ccc
 33 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro
 34 1 5 10
 36 att gtg ggc ctg ggc act tgg aag tct cct ctc ggc aaa gtg aaa gaa
 37 Ile Val Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu
 38 15 20 25 30
 40 gca gtg aag gtg gcc att gat gca gga tat cgg cac att gac tgt gcc
 41 Ala Val Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala
 42 35 40 45
 44 tat gtc tat cag aat gaa cat gaa gtg ggg gaa gcc atc caa gag aag
 45 Tyr Val Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys
 46 50 55 60
 48 atc caa gag aag gct gtg aag cgg gag gac ctg ttc atc gtc agc aag
 49 Ile Gln Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys
 50 65 70 75
 52 ttg tgg ccc act tcc aga tcg aga agc tct tgaacaaacc tggactgaaa
 53 Leu Trp Pro Thr Ser Arg Ser Arg Ser Ser
 54 80 85
 56 tataaaccag tgactaacca ggttgagtgt caccataacc tcacgcagga gaaactgatc
 58 cagtactgcc actccaaggg catcaccgtt acggcctaca gcccctggg ctctccggat
 60 agaccttggg ccaagccaga agacccttcc ctgctggagg atcccaagat taaggagatt
 62 gctgcaaagc acaaaaaaac cgcagcccag gttctgatcc gtttccatat ccagagggaat
 64 gtgattgtca tccccaaagtc tgtgacacca gcacgcattg ttgagaacat tcaggtcttt
 66 gacttttaaat tgagtgtatga ggagatggca accataactca gcttcaacag aaactggagg
 68 gcctgtaacg tgttgcaatc ctctcatttg gaagactatc ccttcgatgc agaattattga

pg 1-3
 Does Not Comply
 Corrected Diskette Needed

insufficient explanation - give source
 of genetic
 material

(see item 11
 on Enov
 summary
 sheet)

RAW SEQUENCE LISTING

DATE: 03/29/2004

PATENT APPLICATION: US/10/653,681

TIME: 13:18:10

Input Set : A:\U 014798-3.ST25.txt

Output Set: N:\CRF4\03292004\J653681.raw

```

70 gggtgaatct cctgggtgaga ttatacagga gattctcttt ctctcgtgaa gtgtgactac      833
72 ctccactcat gtccatttt agccaagctt atttaagatc acagtgaact tagtcctgtt      893
74 atagacgaga atcgagggtgc tgttttagac atttatttct gtatgttcaa ctaggatcag      953
76 aatatcacag aaaagcatgg cttgaataag gaaatgacaa ttttttccac ttatctgac      1013
78 agaacaaatg tttattaagc atcagaaact ctgccaacac tgaggatgta aagatcaata      1073
80 aaaaaataaa taatcat                                          1090
83 <210> SEQ ID NO: 2
84 <211> LENGTH: 88
85 <212> TYPE: PRT
86 <213> ORGANISM: ARTIFICIAL
88 <220> FEATURE:
89 <223> OTHER INFORMATION: UNKNOWN
91 <400> SEQUENCE: 2
93 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro Ile Val
94 1 5 10 15
97 Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu Ala Val
98 20 25 30
101 Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala Tyr Val
102 35 40 45
105 Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys Ile Gln
106 50 55 60
109 Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys Leu Trp
110 65 70 75 80
113 Pro Thr Ser Arg Ser Arg Ser Ser
114 85
117 <210> SEQ ID NO: 3
118 <211> LENGTH: 1279
119 <212> TYPE: DNA
120 <213> ORGANISM: ARTIFICIAL
122 <220> FEATURE:
123 <223> OTHER INFORMATION: UNKNOWN
126 <220> FEATURE:
127 <221> NAME/KEY: CDS
128 <222> LOCATION: (70)..(804)
130 <400> SEQUENCE: 3
131 caaaaacagc aacagaaagc aggacgtgag acttctacct gctcactcag aatcatttct      60
133 gcaccaacc atg gcc acg ttt gtg gag ctc agt acc aaa gcc aag atg ccc      111
134 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro
135 1 5 10
137 att gtg ggc ctg ggc act tgg aag tct cct ctc ggc aaa gtg aaa gaa      159
138 Ile Val Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu
139 15 20 25 30
141 gca gtg aag gtg gcc att gat gca gga tat cgg cac att gac tgt gcc      207
142 Ala Val Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala
143 35 40 45
145 tat gtc tat cag aat gaa cat gaa gtg ggg gaa gcc atc caa gag aag      255
146 Tyr Val Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys
147 50 55 60
149 atc caa gag aag gct gtg aag cgg gag gac ctg ttc atc gtc agc aag      303

```

same error

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/653,681

DATE: 03/29/2004

TIME: 13:18:10

Input Set : A:\U 014798-3.ST25.txt

Output Set: N:\CRF4\03292004\J653681.raw

```

150 Ile Gln Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys
151          65          70          75
153 ttg tgg ccc act ttc ttt gag aga ccc ctt gtg agg aaa gcc ttt gag      351
154 Leu Trp Pro Thr Phe Phe Glu Arg Pro Leu Val Arg Lys Ala Phe Glu
155      80          85          90
157 aag acc ctc aag gac ctg aag ctg agc tat ctg gac gtc tat ctt att      399
158 Lys Thr Leu Lys Asp Leu Lys Leu Ser Tyr Leu Asp Val Tyr Leu Ile
159 95          100          105          110
161 cac tgg cca cag gga ttc aag tct ggg gat gac ctt ttc ccc aaa gat      447
162 His Trp Pro Gln Gly Phe Lys Ser Gly Asp Asp Leu Phe Pro Lys Asp
163          115          120          125
165 gat aaa ggt aat gcc atc ggt gga aaa gca acg ttc ttg gat gcc tgg      495
166 Asp Lys Gly Asn Ala Ile Gly Gly Lys Ala Thr Phe Leu Asp Ala Trp
167          130          135          140
169 gag gcc atg gag gag ctg gtg gat gag ggg ctg gtg aaa gcc ctt ggg      543
170 Glu Ala Met Glu Glu Leu Val Asp Glu Gly Leu Val Lys Ala Leu Gly
171      145          150          155
173 gtc tcc aat ttc agc cac ttc cag atc gag aag ctc ttg aac aaa cct      591
174 Val Ser Asn Phe Ser His Phe Gln Ile Glu Lys Leu Leu Asn Lys Pro
175      160          165          170
177 gga ctg aaa tat aaa cca gtg act aac cag gtt gag tgt cac cca tac      639
178 Gly Leu Lys Tyr Lys Pro Val Thr Asn Gln Val Glu Cys His Pro Tyr
179 175          180          185          190
181 ctc acg cag gag aaa ctg atc cag tac tgc cac tcc aag ggc atc acc      687
182 Leu Thr Gln Glu Lys Leu Ile Gln Tyr Cys His Ser Lys Gly Ile Thr
183          195          200          205
185 gtt acg gcc tac agc ccc ctg ggc tct ccg gat aga cct tgg gcc aag      735
186 Val Thr Ala Tyr Ser Pro Leu Gly Ser Pro Asp Arg Pro Trp Ala Lys
187          210          215          220
189 cca gaa gac cct tcc ctg ctg gag gat ccc aag att aag gag att gct      783
190 Pro Glu Asp Pro Ser Leu Leu Glu Asp Pro Lys Ile Lys Glu Ile Ala
191          225          230          235
193 gca aag cac tcc cca agt ctg tgacaccagc acgcattggtt gagaacattc      834
194 Ala Lys His Ser Pro Ser Leu
195      240          245
197 aggtctttga ctttaaattg agtgatgagg agatggcaac catactcagc ttcaacagaa      894
199 actggagggc ctgtaacgtg ttgcaatcct ctcatttgga agactatccc ttcgatgcag      954
201 aatattgagg ttgaatctcc tgggtgagatt atacaggaga ttctctttct tcgctgaagt      1014
203 gtgactacct ccaactcatgt cccatttttag ccaagcttat ttaagatcac agtgaactta      1074
205 gtcctgttat agacgagaaat cgaggtgctg ttttagacat ttatttctgt atgttcaact      1134
207 aggatcagaa tatcacagaa aagcatggct tgaataagga aatgacaatt tttccactt      1194
209 atctgatcag aacaaatggt tattaagcat cagaaactct gccaacactg aggatgtaaa      1254
211 gatcaataaa aaaaataata atcat      1279
214 <210> SEQ ID NO: 4
215 <211> LENGTH: 245
216 <212> TYPE: PRT
217 <213> ORGANISM: ARTIFICIAL
219 <220> FEATURE:
220 <223> OTHER INFORMATION: UNKNOWN

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/653,681

DATE: 03/29/2004

TIME: 13:18:10

Input Set : A:\U 014798-3.ST25.txt

Output Set: N:\CRF4\03292004\J653681.raw

222 <400> SEQUENCE: 4

```

224 Met Ala Thr Phe Val Glu Leu Ser Thr Lys Ala Lys Met Pro Ile Val
225 1          5          10          15
228 Gly Leu Gly Thr Trp Lys Ser Pro Leu Gly Lys Val Lys Glu Ala Val
229          20          25          30
232 Lys Val Ala Ile Asp Ala Gly Tyr Arg His Ile Asp Cys Ala Tyr Val
233          35          40          45
236 Tyr Gln Asn Glu His Glu Val Gly Glu Ala Ile Gln Glu Lys Ile Gln
237          50          55          60
240 Glu Lys Ala Val Lys Arg Glu Asp Leu Phe Ile Val Ser Lys Leu Trp
241 65          70          75          80
244 Pro Thr Phe Phe Glu Arg Pro Leu Val Arg Lys Ala Phe Glu Lys Thr
245          85          90          95
248 Leu Lys Asp Leu Lys Leu Ser Tyr Leu Asp Val Tyr Leu Ile His Trp
249          100         105         110
252 Pro Gln Gly Phe Lys Ser Gly Asp Asp Leu Phe Pro Lys Asp Asp Lys
253          115         120         125
256 Gly Asn Ala Ile Gly Gly Lys Ala Thr Phe Leu Asp Ala Trp Glu Ala
257          130         135         140
260 Met Glu Glu Leu Val Asp Glu Gly Leu Val Lys Ala Leu Gly Val Ser
261 145         150         155         160
264 Asn Phe Ser His Phe Gln Ile Glu Lys Leu Leu Asn Lys Pro Gly Leu
265          165         170         175
268 Lys Tyr Lys Pro Val Thr Asn Gln Val Glu Cys His Pro Tyr Leu Thr
269          180         185         190
272 Gln Glu Lys Leu Ile Gln Tyr Cys His Ser Lys Gly Ile Thr Val Thr
273          195         200         205
276 Ala Tyr Ser Pro Leu Gly Ser Pro Asp Arg Pro Trp Ala Lys Pro Glu
277          210         215         220
280 Asp Pro Ser Leu Leu Glu Asp Pro Lys Ile Lys Glu Ile Ala Ala Lys
281 225         230         235         240
284 His Ser Pro Ser Leu
285          245

```

RAW SEQUENCE LISTING ERROR SUMMARY DATE: 03/29/2004
PATENT APPLICATION: US/10/653,681 TIME: 13:18:11

Input Set : A:\U 014798-3.ST25.txt
Output Set: N:\CRF4\03292004\J653681.raw

Invalid <213> Response:

Use of "Artificial" only as "<213> Organism" response is incomplete,
per 1.823(b) of New Sequence Rules. Valid response is Artificial Sequence.

Seq#:1,2,3,4

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/653,681

DATE: 03/29/2004

TIME: 13:18:11

Input Set : A:\U 014798-3.ST25.txt

Output Set: N:\CRF4\03292004\J653681.raw